



Human cochlear tuning estimates from stimulus-frequency otoacoustic emissions

Bentsen, Thomas; Harte, James; Dau, Torsten

Published in:
Journal of the Acoustical Society of America

Link to article, DOI:
[10.1121/1.3575596](https://doi.org/10.1121/1.3575596)

Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Bentsen, T., Harte, J., & Dau, T. (2011). Human cochlear tuning estimates from stimulus-frequency otoacoustic emissions. *Journal of the Acoustical Society of America*, 129(6), 3797-3807. <https://doi.org/10.1121/1.3575596>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Human cochlear tuning estimates from stimulus-frequency otoacoustic emissions

Thomas Bentsen, James M. Harte,^{a)} and Torsten Dau

Center for Applied Hearing Research, Department of Electrical Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

(Received 21 August 2010; revised 1 March 2011; accepted 16 March 2011)

Two objective measures of human cochlear tuning, using stimulus-frequency otoacoustic emissions (SFOAE), have been proposed. One measure used SFOAE phase-gradient delay and the other two-tone suppression (2TS) tuning curves. Here, it is hypothesized that the two measures lead to different frequency functions in the same listener. Two experiments were conducted in ten young adult normal-hearing listeners in three frequency bands (1-2 kHz, 3-4 kHz and 5-6 kHz). Experiment 1 recorded SFOAE latency as a function of stimulus frequency, and experiment 2 recorded 2TS iso-input tuning curves. In both cases, the output was converted into a sharpness-of-tuning factor based on the equivalent rectangular bandwidth. In both experiments, sharpness-of-tuning curves were shown to be frequency dependent, yielding sharper relative tuning with increasing frequency. Only a weak frequency dependence of the sharpness-of-tuning curves was observed for experiment 2, consistent with objective and behavioural estimates from the literature. Most importantly, the absolute difference between the two tuning estimates was very large and statistically significant. It is argued that the 2TS estimates of cochlear tuning likely represents the underlying properties of the suppression mechanism, and not necessarily cochlear tuning. Thus the phase-gradient delay estimate is the most likely one to reflect cochlear tuning. © 2011 Acoustical Society of America.

[DOI: 10.1121/1.3575596]

PACS number(s): 43.64.Jb, 43.64.Kc [WPS]

Pages: 3797–3807

I. INTRODUCTION

Cochlear tuning refers to the ability of the cochlea to resolve complex sounds into individual tones. The current “gold standards” for assessing cochlear tuning are behavioral measures, based on forward- or simultaneous-masking procedures and can be regarded as a measure of basilar membrane (BM) filtering (Oxenham and Shera, 2003). It would be desirable to have reliable objective measures to assess human cochlear tuning across frequency and subjects, e.g., for neonates and uncooperative subjects. One such objective method makes use of stimulus-frequency otoacoustic emissions (SFOAEs). Otoacoustic emissions (OAEs) can be considered as a leakage of acoustic energy from the active, frequency-selective transmission mechanism that takes place in the cochlea (Kemp, 2008). SFOAEs are recorded at the same frequency as an applied pure-tone stimulus (Kalluri and Shera, 2007a). SFOAEs are easy-to-record and are appropriate to use to estimate frequency selectivity, as pure tones at moderate levels only excite a narrow region of the BM. Two SFOAE measures of cochlear tuning, based on latency estimation and on two-tone suppression (2TS) tuning curves, have been proposed in the literature (Shera *et al.*, 2002; Shera and Guinan, 2003; Schairer *et al.*, 2006; Keefe *et al.*, 2008; Lineton and Wildgoose, 2009). These two SFOAE measures should both be correlated with human auditory filter bandwidth (Lineton and Wildgoose, 2009) and thus provide estimates on the sharpness of cochlear tuning.

A. SFOAE latency

Using SFOAE latency to derive an estimate of cochlear tuning was originally motivated by predictions from the predominant hypothesis for OAE generation, the coherent reflection filtering theory (Shera and Guinan, 1999; Shera *et al.*, 2002; Shera and Guinan, 2003). This approach to estimate cochlear tuning has been used extensively in the literature (Shera *et al.*, 2002; Shera and Guinan, 2003; Schairer *et al.*, 2006) at low stimuli levels (<40 dB SPL), where the cochlea is assumed to behave linearly. The coherent reflection filtering theory explains the evolution of the SFOAE phase with frequency which depends on basic cochlear mechanical properties. From the SFOAE phase, the SFOAE latency, τ_{SFOAE} , can be calculated as the group delay:

$$\tau_{\text{SFOAE}}(f_p) = -\frac{1}{2\pi} \frac{d\phi(f_p)}{df_p} \quad (1)$$

where f_p is the stimulus frequency (also denoted the probe frequency). The link between SFOAE delay and basilar membrane (BM) delay, usually made using coherent reflection filtering theory, is *not* a requirement for estimating cochlear tuning. As recently clarified in Shera *et al.* (2010) and implicit in the original Shera *et al.* (2002) study, the link between SFOAE latency and cochlear tuning can be made by exploiting a purely empirical relationship. For this purpose, it is useful to express SFOAE latency in dimensionless periods of the characteristic frequency (CF) as

$$N_{\text{SFOAE}} = \tau_{\text{SFOAE}} \cdot f_p. \quad (2)$$

^{a)}Author to whom correspondence should be addressed. Electronic mail: harte_j@wmg.warwick.ac.uk

Cochlear tuning can be evaluated using the sharpness-of-tuning Q -factor. To compare tuning from objective measures with behavioural measures, Q_{ERB} is used and is based on the equivalent rectangular bandwidth (ERB). This is defined as the bandwidth of a corresponding rectangular filter, with the same peak amplitude that passes the same amount of power as the original filter when the rectangular filter is driven by white noise (Shera *et al.*, 2002). Classically, Zweig (1976) demonstrated that the auditory filters show many characteristics of a minimum phase system. Specifically, the bandwidths and slopes of the phase functions are inversely proportional, with steeper slopes (larger N_{SFOAE}) corresponding to smaller bandwidths (larger Q_{ERB}).

A constant of proportionality or tuning ratio, k , was defined as the ratio of tuning sharpness to SFOAE phase gradient delay in periods. Shera *et al.* (2010) argued that the tuning ratio has a species invariant form in cat, guinea pig and chinchilla, found from published auditory nerve (AN) tuning data and SFOAE latency. This approximate species invariance allowed Shera *et al.* (2010) to estimate cochlear tuning in humans from recorded SFOAE delays. Shera *et al.* (2002) derived a similar and related empirical measure based on a power-law fit to guinea pig and cat AN tuning and SFOAE latency data. This work was criticized as the formulation used by Shera *et al.* (2002) was to convert from SFOAE delay to *estimated* BM delay. The estimates were based on simplified coherent reflection filtering theory predictions, which have come under challenge in the literature (Siegel *et al.*, 2005; Ren, 2004). In fact, the Shera *et al.* (2002) procedure does not rely on any relationship between SFOAE and BM delays, *nor* on any model of OAE generation. In the present study, the procedure from Shera *et al.* (2002) is used with the frequency dependent constant of proportionality altered to reflect how tuning is empirically related with SFOAE delay in periods. Human cochlear tuning is estimated in the present study by

$$Q_{\text{ERB}} = k \cdot N_{\text{SFOAE}} \quad (3)$$

where k is the dimensionless tuning ratio, equivalent to the reciprocal of the product of the filter bandwidth and τ_{SFOAE} . In this estimate both k , N_{SFOAE} and hence Q_{ERB} depend on the characteristic frequency (CF). A power-law form of human tuning ratio, k , is given as

$$k = \beta f_p^\alpha \quad (4)$$

where $\alpha = -0.07 \pm 0.06$ and $\beta = 1.15 \pm 0.16$ (half the value given in Shera *et al.* (2002) to remove the reference to *estimated* BM delay). Equations (1)–(3) thus provide an empirically-based relation between SFOAE phase and cochlear tuning, completely independent of any theory of OAE generation. As pointed out by Shera *et al.* (2010), the primary assumption of this method is the species invariance of the tuning ratio.

B. Two-tone suppression

Another estimate of cochlear tuning can be made using 2TS tuning curves (Brass and Kemp, 1993; Keefe *et al.*,

2008; Lineton and Wildgoose, 2009). 2TS refers to the phenomena where one tone, denoted the suppressor tone, reduces the amplitude of the response to the probe tone. 2TS tuning curves can be obtained by fixing the probe tone while sweeping the suppressor tone around the probe frequency. When the suppressor frequency is closer to the probe frequency, a greater degree of suppression is observed. Thus, 2TS tuning curves appear to have a bandpass shape, from which Q_{ERB} can be estimated. Historical studies have linked 2TS tuning curve bandwidth to cochlear tuning (Brass and Kemp, 1993; Keefe *et al.*, 2008; Lineton and Wildgoose, 2009). It has been argued that the 2TS tuning curve bandwidth depends on the amount of suppression, which is a function of the amount of overlap between the probe tone traveling wave (TW) and the suppressor tone TW on the BM, which again is related to frequency selectivity and thus auditory filter bandwidth. As opposed to the empirically-derived relation between SFOAE latency and Q_{ERB} , the link between 2TS tuning curves and cochlear tuning relies on this weaker heuristic argument. A more robust theoretical background is desired but is lacking in the literature, as highlighted by Brass and Kemp (1993), Keefe *et al.* (2008) and Lineton and Wildgoose (2009).

C. Historical tuning estimates and scope of this study

Different objective estimates of Q_{ERB} as a function of frequency have been obtained in the literature. Most studies using SFOAE latency concluded that Q_{ERB} increases with increasing CF, from 0.4 to 10 kHz (Shera *et al.*, 2010, 2002; Shera and Guinan, 2003). This implies that the auditory filters sharpen with frequency. In contrast, however, Schairer *et al.* (2006) found Q_{ERB} to be invariant to frequency even though SFOAE latency was also used. This discrepancy between the Q_{ERB} frequency dependence based on SFOAE latency, particularly at $\text{CF} \geq 4$ kHz, has not been resolved yet (Schairer *et al.*, 2006). Keefe *et al.* (2008) measured 2TS tuning curves in the range of 1–4 kHz and showed Q_{ERB} estimates that were independent of frequency. Thus, different Q_{ERB} frequency functions have been observed in historical studies (Shera *et al.*, 2002, 2010; Schairer *et al.*, 2006; Keefe *et al.*, 2008) across the two different SFOAE measures, and the reason for the differences has remained unclear. Potentially, the two SFOAE measures reflect the tuning properties of the cochlea in different excitation states. In order to further examine this in the present study, both paradigms were implemented in the same subjects in the range from 1–6 kHz, using the same experimental setup. Lineton and Wildgoose's (2009) study provided some indirect evidence, based on numerical cochlear modeling, to suggest that group delay (GD) and two-tone suppression (2TS) methods should provide similar bandwidth estimates. They demonstrated via their model that the tuning measures based on GD and 2TS had a weak correlation. However, they failed to demonstrate a significant correlation in the two derived estimates of tuning across 16 human subjects. Lineton and Wildgoose (2009) limited their analysis to a 1–2 kHz band. By restricting their frequency range, the small correlation predicted from modeling would likely have been so small due to the

inherent variability seen in SFOAE data. In the present study, the frequency range was expanded with two additional bands at 3–4 and 5–6 kHz in an attempt to clarify if the small correlation predicted by Lineton and Wildgoose (2009) exists. Results of the present study were compared to previous studies (Shera *et al.*, 2002, 2010; Schairer *et al.*, 2006; Keefe *et al.*, 2008).

II. METHODS

A. Subjects

A total of ten NH subjects (seven males and three females) participated in the study, their ages ranged from 23–30 years. Five right and five left ears were chosen randomly from the subjects. All subjects showed audiometric thresholds ≤ 15 dB HL. The experiments were conducted in an acoustically shielded audiometric booth (IEC 268-13).

B. Stimuli and data acquisition

SFOAEs were recorded with the so-called suppression method¹ (Brass and Kemp, 1993; Shera and Guinan, 1999; Schairer *et al.*, 2006; Kalluri and Shera, 2007a,b; Keefe *et al.*, 2008; Lineton and Wildgoose, 2009) using a probe tone and a suppressor tone presented simultaneously. An interleaved, two-interval stimuli procedure (Shera and Guinan, 1999; Kalluri and Shera, 2007a,b) was used. The probe tone was played continuously and the suppressor tone was turned on and off in cycles. The probe and suppressor tones were played through separate earphones, to prevent spurious contamination by earphone distortion (Kalluri and Shera, 2007b). The stimulus was split into a probe and a suppressor segment, where each segment consisted of the following six intervals:

$$\text{segment} = \begin{cases} P_1 P_2 P_3 P_4 P_5 P_6 & \text{probe} \\ O_1 O_2 W_3 S_4 S_5 W_6 & \text{suppressor.} \end{cases} \quad (5)$$

The intervals were equally long with $N = 4096$ samples, or $T = 85.3$ ms, with a sampling rate of 48 kHz. Thus, the total segment duration was $6T = 512$ ms. The probe segment consisted of six identical intervals P_1, P_2, \dots, P_6 of a pure tone with frequency f_p . The frequency f_p was adjusted, so that each interval contained an integer number of cycles of the tone. The suppressor segment contained different intervals. The O_1, O_2 intervals contained zeros and the intervals S_4, S_5 consisted of an adjusted pure tone with frequency f_s and zero initial phase. The W_3, W_4 intervals consisted of the suppressor interval, multiplied by half-Blackman windows that ramped the suppressor on and off. This was done to reduce any unwanted transient emission (Shera and Guinan, 1999). The segments were repeated a fixed number of times $J = 30$, which was considered enough to achieve a sufficiently large signal-to-noise ratio (SNR) across the different subjects.

The stimuli were generated in Matlab and played via a soundcard (RME Fireface800). The probe and the suppressor were sent through two channels to a headphone driver (TDT HB 7), and to subjects, through separate insert-earphones (ER-2) connected to an OAE probe (ER-10B). The recorded

ear canal pressure from the OAE probe was amplified by 40 dB (ER-10dB) and band-pass filtered (Rockland model 852) between 0.5 kHz and 12 kHz, and stored digitally on a PC for off-line analysis.

C. Measurement procedures

Two experiments were conducted: experiment 1 recorded SFOAEs as a function of probe frequency and experiment 2 recorded SFOAE 2TS-tuning curves as a function of suppressor frequency. All stimuli were calibrated using a coupler (IEC 711) in dB SPL. All recordings were repeated twice, so test-retest reliability could be evaluated.

1. Experiment 1: SFOAEs as a function of probe frequency

Low-level SFOAEs could lead to discontinuities in the SFOAE phase curve due to poor signal to noise ratio (SNR). A suppressor of $f_s = 1.15 f_p$ was chosen as this has been shown to result in maximum suppression of the SFOAE component in historical studies (Brass and Kemp, 1991, 1993; Keefe *et al.*, 2008; Lineton and Wildgoose, 2009). Thus, when the SFOAE was extracted from the unsuppressed condition, by subtracting the suppressed response (explained in detail in Sec. II D 2), the largest possible SFOAE residual was obtained. This ensured a reasonable SNR. The probe was swept in frequency in three different bands; band 1 from 1–2 kHz (996.09–2003.91 Hz), band 2 from 3–4 kHz (3000–3996.09 Hz) and 5–6 kHz (5003.91–6000 Hz) with the number of linear-spaced frequency components being 87, 86, and 86, respectively. The frequency step was approximately 11.7 Hz, and was deemed to provide a sufficiently high resolution such that the derived phase curve lacked any ambiguities due to unwrapping.

The probe level L_p and suppressor level L_s were chosen as $L_p = 40$ dB SPL and $L_s = 60$ dB SPL, respectively. These levels were chosen, as they produce sufficiently large SFOAE levels (Kalluri and Shera, 2007a).

2. Experiment 2: SFOAE 2TS-tuning curve as a function of suppressor frequency

The probe frequency f_p was fixed while the suppressor frequency was swept around this. The suppressor frequency $f_s = f_p$ was omitted from the recordings. The range of f_s was chosen based on preliminary recordings in a selected subject to ensure a sufficient range of the 2TS tuning curve. A range from $0.9 f_p$ to $1.6 f_p$ was found appropriate to capture most of the energy below the 2TS tuning curves in each band.²

SFOAE levels typically show a pattern of local maxima and minima across f_p , known as the SFOAE fine structure (Brass and Kemp, 1993; Shera and Guinan, 1999; Kalluri and Shera, 2007a,b). Frequency f_p was chosen for each subject individually at a local maximum in the SFOAE level curves from experiment 1 to produce high-level 2TS tuning curves with a well-defined bandpass shape.

An iso-input recording paradigm was applied, keeping the suppressor at a constant level and varying the degree of suppression of the SFOAE by varying the suppressor

frequency f_s . Identical probe and suppressor levels, $L_p = L_s = 40$ dB SPL, were used. In theory, this should evoke suppressor traveling waves (TWs) along the BM that were similar to the TW of the probe during experiment 1 (Lineton and Wildgoose, 2009).

D. Off-line processing

1. Artifact rejection, averaging and Fourier analysis

Artifact rejection was performed to remove any recording epochs containing large amounts of unwanted noise. First, the initial two repetitions were discarded from the raw signal, to avoid any unwanted transient responses from the transducers (Shera and Guinan, 1999). Secondly, 10% of the noisiest repetitions were discarded in each recorded buffer. Time-domain averaging (Shera and Guinan, 1999; Kalluri and Shera, 2007a,b) was applied, where the probe intervals (X_1, X_2) and probe + suppressor intervals (X_4, X_5) were averaged across repetitions. Finally, (X_1, X_2) and (X_4, X_5) were averaged into one probe and probe + suppressor interval (each interval containing 4096 samples), respectively. The probe and probe + suppressor interval spectra were obtained via the fast Fourier transform.

2. Extracting emission components

During experiment 1, the recorded ear canal pressure can be regarded as the sum of two complex pressure components (Kalluri and Shera, 2007b):

$$p_p(f_p) = p_{p+s:MS}(f_p) + p_{\text{SFOAE}}(f_p), \quad (6)$$

where $p_p(f_p)$ is the complex ear canal pressure at f_p , extracted from the complex spectrum of the probe interval. The second term, $p_{p+s:MS}(f_p)$, is the complex ear canal pressure extracted from the probe + suppressor interval, under the assumption of maximal suppression of SFOAEs. This means that the energy in $p_{p+s:MS}(f_p)$ should correspond to the complex stimulus pressure only, as the SFOAE component has been completely suppressed. The last term in Eq. (6) is the complex SFOAE pressure wave. Rearranging Eq. (6) leads to an estimate of p_{SFOAE} :

$$p_{\text{SFOAE}}(f_p) = [p_p(f_p) - p_{p+s:MS}(f_p)] e^{-2\pi j \tau_{\text{eq}} f_p} \quad (7)$$

The exponential factor in Eq. (7) compensates for the internal recording rig delay, τ_{eq} . The SFOAE level, in dB SPL, at a specific f_p was then calculated from $p_{\text{SFOAE}}(f_p)$:

$$L_{\text{SFOAE}}(f_p) = 20 \log_{10} \left(\frac{|p_{\text{SFOAE}}(f_p)|}{p_{\text{ref}}} \right), \quad (8)$$

where $p_{\text{ref}} = 20 \mu\text{Pa}$. The SFOAE phase was given by

$$\phi(f_p) = \arctan \left\{ \frac{\text{Im}[p_{\text{SFOAE}}(f_p)]}{\text{Re}[p_{\text{SFOAE}}(f_p)]} \right\}, \quad (9)$$

where Eq. (9) returns the wrapped phase in radians. An unwrapping procedure was applied and the phase angles

were converted into cycles of f_p . In each band, the slope of the phase was calculated using the three-point finite difference method (Press *et al.*, 1986) and converted into latency using Eq. (1). Group delay estimates based on phase-gradient are known to be sensitive to noise. Therefore, each latency estimate was weighted by the normalized squared pressure amplitude of p_{SFOAE} (Lineton and Wildgoose, 2009):

$$w(f_p) = \frac{|p_{\text{SFOAE}}(f_p)|^2}{p_{\text{SFOAE,av}}^2}, \quad (10)$$

where $p_{\text{SFOAE,av}}^2$ is an average value of the squared pressure amplitude across the band from $f_{p,1}$ to $f_{p,2}$:

$$p_{\text{SFOAE,av}}^2 = \frac{1}{f_{p,2} - f_{p,1}} \int_{f_{p,1}}^{f_{p,2}} |p_{\text{SFOAE}}(f_p)|^2 df_p. \quad (11)$$

The weighting function w provides more weight to frequencies with correspondingly large emission amplitudes, and penalizes frequencies with low emission amplitude, as these create discontinuities in the unwrapped phase. The average latency across a specific band is given by (Lineton and Wildgoose, 2009)

$$\tau_{\text{SFOAE}} = \frac{1}{f_{p,2} - f_{p,1}} \int_{f_{p,1}}^{f_{p,2}} \tau(f_p) w(f_p) df_p. \quad (12)$$

This leads to three averaged SFOAE latencies, centered in the three bands analyzed.³

In experiment 2, a complex pressure component $p_{2\text{TS}}$, was extracted from the recorded ear canal pressure, using the following complex subtraction:

$$p_{2\text{TS}}(f_p) = [p_p(f_p) - p_{p+s:PS}(f_p)] e^{-2\pi j \tau_{\text{eq}} f_p}. \quad (13)$$

Here, $p_{p+s:PS}(f_p)$ refers to the complex ear canal pressure extracted from the probe + suppressor, and assumes that the SFOAE is only partially suppressed. The degree of suppression observed is a function of suppressor level and frequency. Thus, the resulting complex subtraction in Eq. (13) indicates that $p_{2\text{TS}}(f_p)$ can be interpreted as the complex pressure of the partial SFOAE. For each f_s , $p_{2\text{TS}}(f_p)$ was extracted and converted into a level in dB SPL. Plotting f_s against this level yielded the 2TS iso-input tuning curve.

3. Estimates of the noise floor

The noise floors in experiments 1 and 2 were assessed by calculating an average of adjacent frequency bins to the f_p bin. This method relied on the assumption that the noise of the neighboring frequency bins provided estimates of the noise at the f_p bin, thus assuming that the additive noise was white and stationary. Ten spectral bins to either side of the f_p bin were extracted from the probe and probe + suppressor spectra. A complex subtraction between corresponding noise bins in the probe and probe + suppressor spectra was then carried out to provide estimates of SFOAE noise at each noise bin. The bins were then averaged and converted to dB

SPL. In experiment 2, the suppressor frequency, f_s , was always excluded from the noise estimation procedure.

4. Sharpness-of-tuning calculations

The SFOAE latency estimates from experiment 1 were converted into sharpness-of-tuning factors, Q_{ERB} , using Eqs. (2)–(4). Q_{ERB} was evaluated using center frequencies $\text{CF} = 1.5, 3.5, 5.5$ kHz from each band to provide an average estimate in each of the three bands. Shera *et al.* (2002) did provide fitted means and 95% confidence interval limits of k , valid at $\text{CF} \geq 1$ kHz, corresponding to the basal 60% of the cochlea. Only the fitted mean k function was used here (altered to remove reference to BM delay) to estimate the Q_{ERB} in Eq. (3). Thus, the variation in k , indicated by 95% confidence interval limits of the mean k was not considered here.

As in Keefe *et al.* (2008), the Q_{ERB} was estimated from 2TS tuning curves by squaring the amplitude of the 2TS tuning curves. The area below this curve was then interpreted as the power passed by the 2TS tuning curve. This was integrated over the range of suppressor frequencies $f_{s,1}$ to $f_{s,2}$, as

$$P = \int_{f_{s,1}}^{f_{s,2}} |p_{2\text{TS}}(f_s)|^2 df_s. \quad (14)$$

The ERB was estimated as

$$\text{ERB} = \frac{P}{\max |p_{2\text{TS}}(f_s)|^2}, \quad (15)$$

reflecting the ratio between the total power P , passed by the 2TS tuning curve, and the peak squared amplitude of the 2TS tuning curve. The Q_{ERB} was computed as

$$Q_{\text{ERB}} = \frac{f_p}{\text{ERB}}. \quad (16)$$

III. RESULTS

A. Experiment 1: SFOAEs as a function of probe frequency

Figure 1 shows SFOAE level and phase versus probe frequency, f_p , for one illustrative subject “sr,” recorded in

the three bands 1–3. The top panels show SFOAE level (solid lines) and the noise floor (dashed lines) in dB SPL and the bottom panels show unwrapped SFOAE phase in cycles of 2π . Test-retest recordings (black and gray curves) are closely overlapping each other in all bands, indicating good and robust estimates. SFOAE levels vary between -20 dB to 20 dB SPL across the three bands, showing the characteristic fine structure of local maxima and minima (Brass and Kemp, 1993; Kalluri and Shera, 2007a,b). In all bands, the SNR was generally excellent (20 – 40 dB). The phase curves showed steady, negative slopes in all bands, indicating a delay relative to the stimulus. The phase roll-off stretched over 12 cycles, 7 cycles, and 6 cycles in each respective band. Thus, the slope and the corresponding BM latency decreased across bands, as stated in Eq. (1). The estimated phase, as reported in Eq. (9), was very sensitive to low SNR levels, in particular to low SFOAE levels. This was observed as discontinuities in the unwrapped phase which is evidently seen in band 3 around $f_p = 5.9$ kHz. The weighting function in Eq. (10) reduced the phase contribution in such regions, making the latency estimates robust against noise. Bands showing particularly low SFOAE levels and thus an unstable phase curve were discarded from the sharpness-of-tuning processing (3 of 30 bands).

B. Experiment 2: SFOAE 2TS-tuning curve as a function of suppressor frequency

Figure 2 shows the 2TS iso-input tuning curves (solid lines) and the noise floor (dashed lines), in dB SPL, and the unwrapped phase in cycles as a function of the normalized frequency f_s/f_p for the illustrative subject “sr” in all three bands. The 2TS tuning curves showed a bandpass-shape with levels decreasing across bands. The peak level was typically located just below $f_s/f_p = 1.2$, although there was significant subject variation. Larger SFOAE levels were obtained in experiment 1 than in experiment 2 at similar frequencies. This was due to the larger suppressor level used in experiment 1 ($L_s = 60$ dB SPL) compared to experiment 2 ($L_s = 40$ dB SPL). Test-retest recordings were repeatable in all bands. However, for illustrative purposes, band 2 for subject “sr” is shown where there is not an excellent agreement. The magnitude and phase plots both appear to be offset from

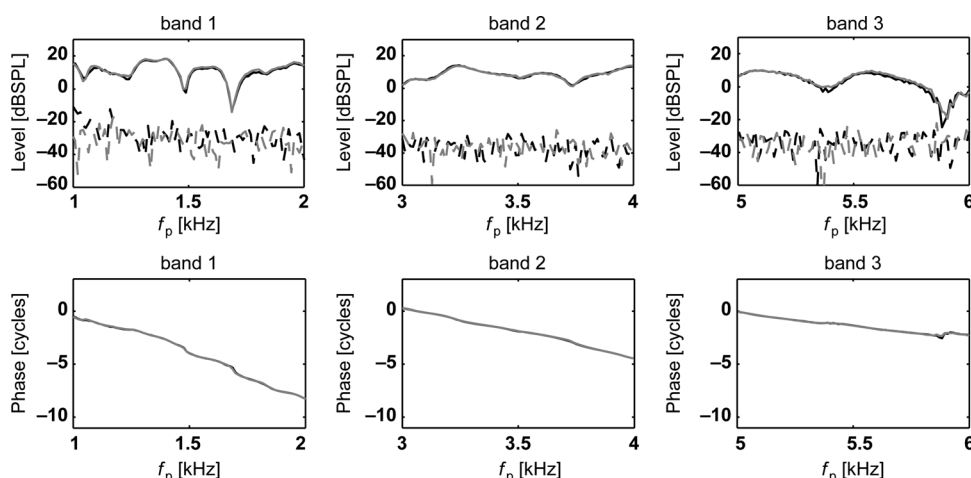


FIG. 1. SFOAE level and phase from subject “sr.” SFOAEs were recorded in the three bands 1–3. Test (black curves) - retest (gray curves) recordings are both shown. Top figures show SFOAE level (solid curves) and the noise floor (dashed curves) in dB SPL versus f_p . Bottom figures show the unwrapped SFOAE phase in cycles of 2π versus f_p .

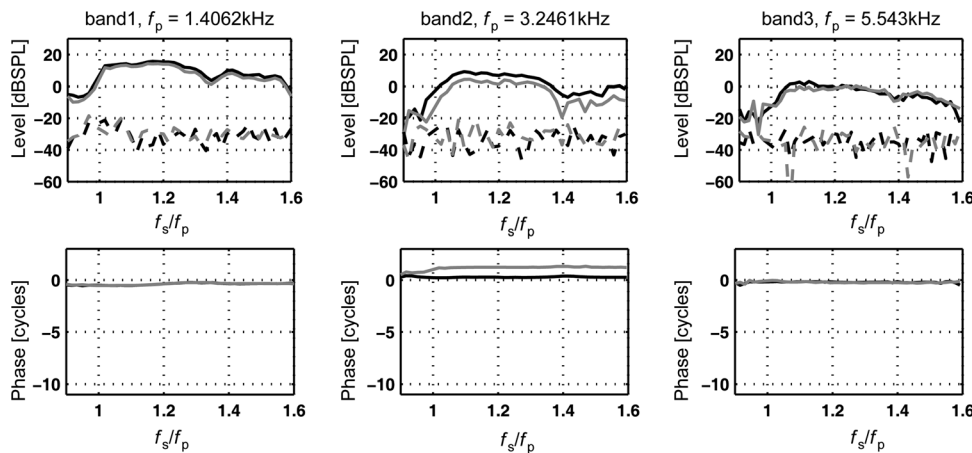


FIG. 2. 2TS iso-input tuning curves and unwrapped phase in cycles from subject “sr.” Curves were recorded in the three bands 1–3 with the choice of f_p labeled on each figure. Test (black curves) – retest (gray curves) recordings are both shown. Top figures show 2TS tuning curves (solid curves) and the noise floor (dashed curves), in dB SPL. Bottom figures show the unwrapped SFOAE phase in cycles of 2π . The frequency axis have been normalized with respect to the specific f_p .

each other. This is likely due to a different probe-insertion depth, which can lead to different excitation levels in the ear canal. This resulted in a constant change in OAE level, however, the phase offset was due to the unwrapping procedure used here. All phase is given relative to the lowest f_s/f_p tested and as can be seen, this is dominated by noise in this band, i.e., random phase. The variation of phase with f_s/f_p was found to be essentially flat in the region of interest, as expected. In general, low SFOAE levels gave rise to poor bandpass shapes and were discarded from the sharpness-of-tuning processing (9 of 30 bands). For accepted subjects, the general tendency was a broadening of level curves across bands 1–3 on the relative frequency scale.

C. Sharpness-of-tuning estimates

Figure 3 shows sharpness-of-tuning Q_{ERB} estimates across CF. The figure shows group mean and group standard error limits of Q_{ERB} estimates for experiments 1 (solid curve with squares) and experiment 2 (solid curve with triangles pointing downwards). Q_{ERB} estimates were averaged across

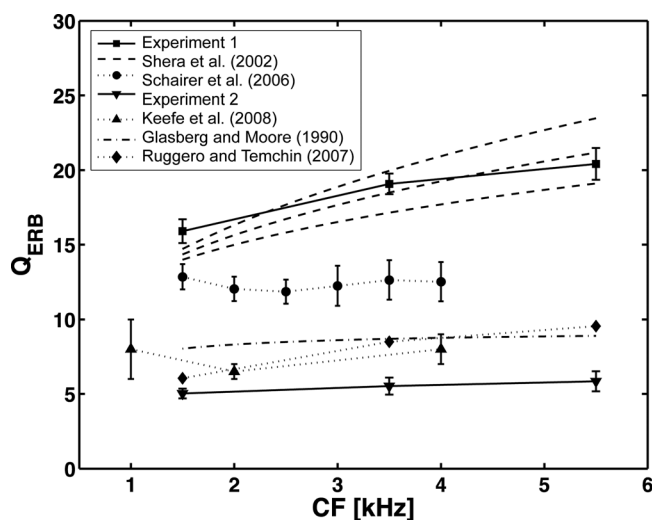


FIG. 3. Means and standard errors of Q_{ERB} measured in experiment 1 (solid line with squares) and 2 (solid line with downwards pointing triangles). Estimated Q_{ERB} curves from historical studies are also shown, including Shera *et al.* (2002) (dashed lines), Schairer *et al.* (2006) (dotted line with circles), Keefe *et al.* (2008) (dotted lines with triangles pointing upwards) and Glasberg and Moore (1990) (dot-dashed line).

test-retest and across subjects. An increase in Q_{ERB} across CF is seen in the group mean curve from experiment 1 (Q_{ERB} values from 16 to 20). The group mean curve from experiment 2 is lower than the curve from experiment 1 and shows a slight increase (Q_{ERB} values from 5 to 6).

A two-way analysis of variance (ANOVA) was applied independently to the Q_{ERB} group data from experiments 1 and 2. Data from experiment 1 showed a significant frequency dependence ($F = 7.4$; $p = 0.006$) while the data from experiment 2 showed a less powerful though still present frequency dependence ($F = 4.8$; $p = 0.04$). In order to investigate the covariation between the data from the two experiments in greater detail, an analysis of covariance (ANOCOVA) was then applied. The ANOCOVA assumes a linear model of each data set. According to Neely *et al.* (1988), for a fixed level of excitation, the basilar membrane latency τ_{BM} can be modelled as a decreasing power law function of the probe frequency f_p as

$$\tau_{\text{BM}} = \alpha f_p^{-b}, \quad (17)$$

where a and b are constants. Assuming that the SFOAE delay, τ_{SFOAE} , is proportional to the BM delay, τ_{BM} , via a constant (Shera and Guinan, 2003; Moleti and Sisto, 2008; Harte *et al.*, 2009) and inserting Eq. (17) in Eqs. (2), (3), and (4) yielded a power-law model of Q_{ERB} as a function of CF. On a double logarithmic scale this implied a linear model of the form

$$\log_{10}[Q_{\text{ERB}}] = \kappa \cdot \log_{10}[f_p] + C, \quad (18)$$

where κ is the slope and C the linear model intercept. Cochlear model simulations by Lineton and Wildgoose (2009) showed a degree of correlation between the two SFOAE measures. For this reason, it is possible to argue that the log-transformed data from experiment 2 should also follow a linear model on the log transformed axes. The ANOCOVA demonstrated fitted slopes which were not statistically different between experiments 1 and 2 (experiment-CF interaction term: $F = 0.82$; $p = 0.37$) at a significance level of 5%. The slope estimate κ (experiment 1: $\hat{\kappa} = 0.198$; $SE_{\hat{\kappa}} = 0.054$; experiment 2: $\hat{\kappa} = 0.100$; $SE_{\hat{\kappa}} = 0.054$) was closer to zero for experiment 2, indicating a slightly weaker frequency

dependence, as previously demonstrated using the two independent ANOVAs. The interception C (experiment 1: $\hat{C} = 1.16$; $SE_{\hat{C}} = 0.027$; experiment 2: $\hat{C} = 0.675$; $SE_{\hat{C}} = 0.027$) differed significantly, implying that the absolute difference of Q_{ERB} was significant between the two estimates.

Both experiments in the present study demonstrated sharpness-of-tuning estimates with significant frequency dependence. As seen in Fig. 3, both indicated sharper cochlear tuning as frequency increased. The ANOCOVA test confirmed a stronger frequency dependence in experiment 1, while the dependence was borderline in experiment 2. Most importantly the absolute differences between the two estimates for Q_{ERB} were very large and significant.

D. Objective sharpness-of-tuning estimates

In addition to the results of the present study, Fig. 3 includes other key objective and subjectively derived Q_{ERB} estimates from the literature. Similar to experiment 1 of the present study, Shera *et al.* (2002) (dashed lines, including confidence intervals) and Schairer *et al.* (2006) (dotted line with filled circles) used SFOAE latency to obtain Q_{ERB} . Estimates of Q_{ERB} from Shera *et al.* (2010) have not been included as they are very similar to Shera *et al.* (2002). As in experiment 2, Keefe *et al.* (2008) (dotted line with filled upward-pointing triangles) used SFOAE 2TS tuning curves. All of these objective studies used a probe level of $L_p = 40$ dB SPL, as in the present study. The Q_{ERB} group mean curve and 95% confidence intervals from Shera *et al.* (2002) (dashed lines) were derived by fitting Q_{ERB} to a power-law function for high frequencies, $CF \geq 1$ kHz. This corresponded to the basal 60% of the cochlea, where Shera *et al.* (2002) argued that SFOAE latencies were verified to be similar to latencies obtained from invasive measurements in laboratory animals.

The group mean curve of experiment 1 in the present study followed the trend of the group mean curve from Shera *et al.* (2002) and those published in Shera *et al.* (2010). At 3.5 kHz and 5.5 kHz, the group mean Q_{ERB} values were within the 95% confidence limits of those estimated by Shera *et al.* (2002). At 1.5 kHz, the group mean Q_{ERB} value was about 1.1 times the estimates from Shera *et al.* (2002).

The group mean Q_{ERB} estimates from experiment 2 (see Fig. 3), demonstrated weaker frequency dependence than for experiment 1. Keefe *et al.* (2008) estimated Q_{ERB} using 2TS iso-suppression tuning curves at frequencies 1, 2, and 4 kHz, and did not demonstrate a significant frequency dependency. However, the standard error of the group estimate of Keefe *et al.* (2008) at 1 kHz was relatively large compared to the standard errors at other frequencies. This would have clouded the test for frequency dependence. The standard errors in the present study are considerably smaller.

IV. DISCUSSION

A. Experiment 1: SFOAE delay estimates for Q_{ERB}

The estimates of Q_{ERB} from experiment 1 were made by assuming a species invariant empirically determined ratio

between cochlear tuning and SFOAE delay in periods. A good agreement with Shera *et al.* (2002) and Shera *et al.*'s (2010) estimates of cochlear tuning in humans were found. One potential source of error in the present study is its assumption that all the recorded SFOAEs arose due to a single SFOAE component with a source restricted to a narrow region of the BM, and therefore reflect tuning properties at that place. In fact, Shera and Guinan (2008) and Choi *et al.* (2008), in chinchilla and humans, respectively, argued that there is strong evidence for interference between two SFOAE components and not only the one predicted by coherent reflection filtering theory. This would manifest in the SFOAE spectra as quasi-periodic notches, or interference patterns between the two components which have similar amplitudes but different phase-gradient delays. Shera and Guinan (2008) and Choi *et al.* (2008) separated these two sources using related signal processing approaches previously reported (Guinan *et al.*, 2003; Kalluri and Shera, 2001; Kim *et al.*, 2001; Knight and Kemp, 2001). The approximate unmixing of these two components allowed the long-latency component to be associated with the SFOAE group delay predicted from coherent reflection filtering theory. The source of the short-latency component is still ambiguous, potentially arising from reflections from more basal locations, or from some distortion or nonlinear source. In practical terms, the addition of the short-latency source has the effect of reducing the phase-gradient delay of the total SFOAE. In the present study, no attempt has been made to compensate for this additional source, as it is believed only to have a significant impact at lower frequencies, $f_p \leq 1$ kHz (Shera and Guinan, 2008). The consequence (if any) of not carrying out this source separation here would be that the SFOAE group delay in experiment 1 may be underestimated. This would result in an increase of Q_{ERB} and would further magnify the differences between SFOAE phase-gradient delay and 2TS estimates. Additionally, none of the other studies considered for comparison with the present study used the source unmixing procedure.

Due to the similar recording methods, Schairer *et al.*'s (2006) Q_{ERB} estimates, shown in Fig. 3, should be very similar to Shera *et al.* (2002) and experiment 1 in the present study. However, Schairer *et al.* (2006) demonstrated that Q_{ERB} was almost independent of CF. The mean Q_{ERB} curves from experiment 1 and from Shera *et al.* (2002) were closest to the mean curve from Schairer *et al.* (2006) at 1.5 kHz, but the difference in values increased with CF. The largest difference occurred at 4 kHz, where Q_{ERB} from Shera *et al.* (2002) and experiment 1 were considerably larger than the values from Schairer *et al.* (2006) by a factor of ≈ 1.5 . This discrepancy in Q_{ERB} trend was commented upon by Schairer *et al.* (2006) and Keefe *et al.* (2008) and is still unresolved. A reasonable first step to resolve this would be to investigate more closely the specific methodology used by Schairer *et al.* (2006) to record/obtain their OAEs or Q_{ERB} post processing. Schairer *et al.* (2006) used the same suppression method (Kalluri and Shera, 2007a) as experiment 1 in the present study and Shera *et al.* (2002), to obtain their SFOAEs. Figures 3 and 4 of Schairer *et al.* (2006) showed examples of phase-versus-frequency which were similar to

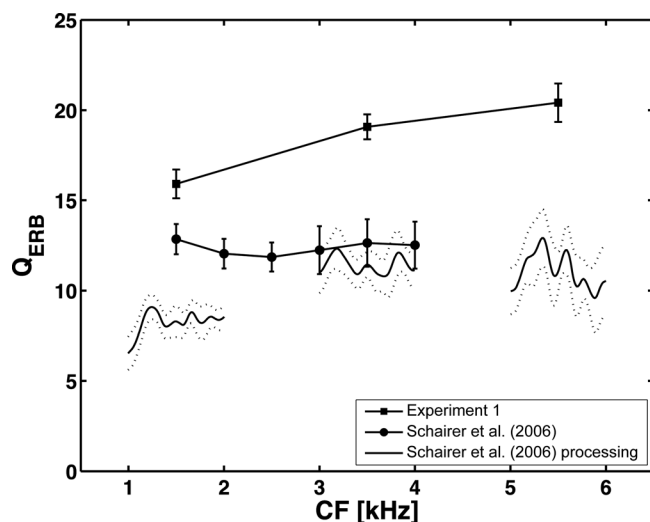


FIG. 4. The group Q_{ERB} mean and SE limits, calculated using the Q_{ERB} post-processing method from Schairer *et al.* (2006) on data from experiment 1. Also, the original Schairer *et al.* (2006) data from Fig. 3 is re-plotted. Refer to Schairer *et al.* (2006) for details on the Q_{ERB} post-processing.

those seen in Figure 1 of the present study. Therefore, basic OAE recording cannot be the source of the Q_{ERB} differences. Thus, the discussion here is restricted to Q_{ERB} post-processing. The Q_{ERB} post-processing steps of Schairer *et al.* (2006) were applied to the SFOAE phase curves from experiment 1. First, the SFOAE phase was smoothed using cubic-spline interpolation across frequency, within each band. Second, the group-delay estimate was calculated, according to Eq. (1), by means of the derivative of the smoothed phase curve. Finally, the group delay estimates were converted into Q_{ERB} estimates (Schairer *et al.*, 2006). The smoothing was performed with each probe frequency weighted by the SNR to make the latency calculation more robust to noise. Figure 4 shows a re-plot of the Q_{ERB} curves from experiment 1 and from Schairer *et al.* (2006). Also, the group mean and standard error limits of Q_{ERB} , derived using Schairer *et al.* (2006) processing steps, are shown. It can be seen that the processing methodology of Schairer *et al.* (2006) yields lower Q_{ERB} values, relative to those employed in experiment 1 here. This highlights the sensitivity of the group delay estimates (and hence Q_{ERB}) to post-processing. However, it should be noted that Shera *et al.* (2002) obtained similar estimates to those of experiment 1 using different post-processing from the present study. The similarity of Q_{ERB} estimates from these two independent methods gives weight to the argument that the phase-gradient delays have been correctly obtained here.

B. Experiment 2: SFOAE 2TS estimates for Q_{ERB}

Results from experiment 2 yielded smaller Q_{ERB} values than previous estimates by Keefe *et al.* (2008), by a factor of 0.7. These estimates were derived using two different paradigms: the iso-input paradigm (Lineton and Wildgoose, 2009) was applied in the present study and the iso-suppression paradigm (Kemp and Chum, 1980; Brass and Kemp, 1993; Keefe *et al.*, 2008) was applied in Keefe *et al.* (2008). In the latter paradigm, the amount of suppression is kept fixed and the 2TS tuning curve is defined as the set of L_s val-

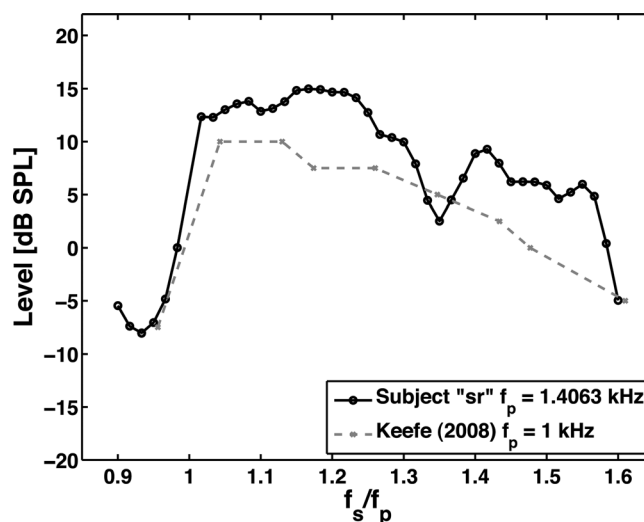


FIG. 5. Comparing the 2TS iso-input and iso-suppression recording paradigm. A 2TS tuning curve recorded in “band 1” on subject “sr” is shown together with data points from Figure 1 Keefe *et al.* (2008). The data points consist of read-off values of the SFOAE residual, recorded using $f_p = 1$ kHz at frequency ratios f_s/f_p at 0.260, 0.348, 0.522, 0.609, 0.695, 0.825, 0.956, 1.043, 1.130, 1.174, 1.260, 1.347, 1.434, 1.477, and 1.61. The data points were read-off at levels $L_p = 40$ dB SPL and $L_s = 40$ dB SPL.

ues, producing the same SFOAE residual. Keefe *et al.* (2008) recorded these SFOAE residuals against different L_s for selected subjects at $f_p = 1$ kHz at specific f_s/f_p ratios and transformed them into subject-specific 2TS iso-suppression tuning curves. The SFOAE residual curves could, however, also generate estimates of the subject-specific iso-input tuning curves by using data points at identical L_p and L_s values. Figure 5 compares a 2TS iso-input tuning curve, estimated using SFOAE residual curves from Keefe *et al.* (2008), to the recorded iso-input tuning curve in subject “sr” for the present study. No obvious, qualitative difference in bandpass shape is seen and thus, the recording paradigm is not expected to explain the Q_{ERB} calculations remarkably.

C. Covariation of GD and 2TS estimates for Q_{ERB}

Lineton and Wildgoose (2009) predicted, via numerical modelling, a weak correlation between auditory filter bandwidth estimated from SFOAE GD and 2TS. Lineton and Wildgoose (2009) found no statistically significant correlation between experimental estimates of BW from GD and 2TS in their 16 adult human test subjects. In the present study, Q_{ERB} rather than auditory filter BW was investigated though they are inversely proportional to each other and hence should also be correlated.

It was hypothesised here that the individual subject variability in SFOAE GD and 2TS curve estimates that arise due to the variability in auditory filter BW and/or Q_{ERB} is lower than the implicit variability due to individual OAE generation mechanisms. By simply looking over a broader frequency range, it was hoped here that the covariation between the two estimates of Q_{ERB} would surface. The results from the ANCOVA demonstrated significant covariation between Q_{ERB} estimated from GD and 2TS across frequency. This leads to the conclusion that the two measures are correlated. However, there is a significant absolute difference between the measures.

D. Comparison with behavioral sharpness-of-tuning estimates

Based on the Q_{ERB} estimates of the present study, one could argue that the two SFOAE measures represent the tuning properties of the cochlear in different excitation states. Behavioral studies of sharpness-of-tuning were also included in Fig. 3, based on forward masking (Shera *et al.*, 2002) (dashed lines) and simultaneous masking (Glasberg and Moore, 1990) (dot-dashed lines) procedures. The Q_{ERB} from Shera *et al.* (2002) was obtained by averaging human SFOAE and forward masking data. The similarities of the Q_{ERB} estimates of Shera *et al.* (2002) with experiment 1 of the present study were discussed above. The simultaneous masking estimates of tuning were illustrated by a polynomial fit based on the notched-noise masking paradigm Glasberg and Moore (1990) and using the concept of equivalent rectangular bandwidth (ERB):

$$\text{ERB} = 24.7(4.37\text{CF}/1000 + 1). \quad (19)$$

Combining this with Eq. (16) yielded Q_{ERB} estimates with a weak frequency dependence (see Fig. 3). The observed Q_{ERB} trend was qualitatively similar to 2TS-derived Q_{ERB} curves from experiment 2 and from Keefe *et al.* (2008).

Simultaneous masking procedures include unavoidable suppression between the signal and the masker (Keefe *et al.*, 2008; Rodriguez *et al.*, 2010), while non-simultaneous masking is not affected by suppression. The striking similarity of the simultaneous masking Q_{ERB} estimates to the SFOAE 2TS-derived estimates suggests that suppression similarly affects both measures. It is argued here that suppression does not affect the estimates of tuning from experiment 1. In this experiment, a second suppressor tone is presented in order to totally suppress the SFOAE (by using an optimal suppressor frequency $f_s = 1.15 f_p$ and a higher level $L_s = 60$ dB SPL). The total SFOAE pressure is indirectly inferred from this completely suppressed recording containing only the stimulus pressure (Guinan *et al.*, 2003; Schairer *et al.*, 2003; Schairer and Keefe, 2005). In experiment 1, the SFOAE phase as a function of excitation frequency is of interest. In experiment 2, the suppressor level is reduced to 40 dB and results in a partially suppressed SFOAE whose magnitude is a function of suppressor frequency.

Keefe *et al.* (2008) have shown that the phase of SFOAEs recorded using the suppressor method is almost independent of the suppressor frequency. Thus, in experiment 2, where the suppressor frequency is varied one would expect the phase to be flat, as can be seen in Fig. 2. It is argued here that the accumulated phase is due to the travel time to the characteristic place of the probe frequency peak of the TW, both for experiment 1 and 2. This is strong evidence to suggest that the place of generation of SFOAE does not vary as suppressor frequency varies. This argues that the SFOAE residual is generated on the BM in the region of tonotopic place, regardless of whether the suppression method is used to record SFOAE. This was supported by Kalluri and Shera (2007a), who demonstrated that SFOAEs have approximately the same amplitude and phase when recorded from methods

that do not make use of a second tone (e.g., compression or spectral smoothing). Therefore, the results of experiment 1 are assumed not to be affected by the suppression mechanism. In experiment 2, 2TS alters the magnitude of the SFOAE as a function of suppressor frequency, even though the recorded emission is generated from the characteristic place of the probe frequency. Therefore, experiment 2 may represent the underlying mechanics of the suppression mechanism, and not necessarily cochlear tuning.

E. Objective sharpness-of-tuning estimates based on BM delay

In experiment 1, the estimates of Q_{ERB} rely on an empirical relationship between cochlear tuning and SFOAE delay. Ruggero and Temchin (2007) offered a novel estimate of *in vivo* BM group delay in humans using post-mortem delay estimates, with the post-mortem effects compensated for via an empirical relation derived from experimental animals. Namely, Ruggero and Temchin (2007) assumed that the quantitative relationship between post-mortem and *in vivo* BM group delay is approximately the same across all mammalian species. Thus, the post-mortem human BM delay could be used to predict *in vivo* delay. From Fig. 7 from Ruggero and Temchin (2007), BM group delay was 1.8, 1.15, and 0.85 ms at 1.5, 3.5, and 5.5 kHz, respectively. In order to convert Ruggero and Temchin's (2007) GD estimates to cochlear tuning estimates comparable with the present study, one must assume that SFOAE GD correctly reflects BM delay *in vivo*. This assumption remains to be tested however. If one assumes a model of SFOAE generation where BM group delay and SFOAE delay are related by a frequency independent constant (generally accepted in the basal portion of the cochlea), then Q_{ERB} estimates based on Ruggero and Temchin (2007) can be made. It is initially assumed that this constant of proportionality is 2, based on simplified coherent reflection filtering theory (Shera and Guinan, 2003). Fig. 3 (dotted line with filled diamonds) shows the Q_{ERB} estimates made here based on Ruggero and Temchin's (2007) human BM delay. Much shorter delays, than obtained by SFOAE phase-gradient methods, were observed by Ruggero and Temchin (2007). This led to Q_{ERB} estimates closer to the 2TS estimates of the present study. Ruggero and Temchin's (2007) estimated BM group delays from the post-mortem delays are of the order of a factor of 5 smaller than SFOAE delays recorded here. There is no supporting *in vivo* evidence in experimental animals to suggest such a factor difference, whereas the evidence points to a factor of nearly 2 (Shera *et al.*, 2002; Shera and Guinan, 2003; Siegel *et al.*, 2005; Moleti and Sisto, 2008; Harte *et al.*, 2009) or closer to 1 at low frequencies (Siegel *et al.*, 2005; Ren, 2004) if one does not account for both long and short latency SFOAE components (Shera and Guinan, 2008; Choi *et al.*, 2008).

F. Influence of stimulus level on estimates of Q_{ERB}

The present study estimated cochlear tuning at a single excitation level of 40 dB SPL. However, it should be commented that cochlear tuning is a level-dependent property, where physiological evidence in animals points to broadening

BM excitation patterns for increasing excitation level (e.g., [Ruggero et al., 1997](#)). Indirect estimates based on BM group delay estimates from auditory brainstem responses to tone bursts would tend to agree with this ([Neely et al., 1988](#)). [Schairer et al. \(2006\)](#) demonstrated that Q_{ERB} estimated from SFOAE group delay decreased as excitation level increased from 40 to 70 dB SPL, as one would expect the SFOAE group delay to decrease (BW increase as excitation patterns broaden) with excitation level. The estimates of [Keefe et al. \(2008\)](#) of Q_{ERB} based on 2TS demonstrated no significant level dependence as excitation level was increased from 30 to 60 dB SPL. Future studies looking at higher excitation levels should take care to see if there is any influence of a second SFOAE component on the Q_{ERB} , as mentioned in Sec. IV A. Using the source unmixing procedure of [Shera and Guinan \(2008\)](#) and [Choi et al. \(2008\)](#) may be necessary.

G. Suppression and cochlear tuning

Two-tone suppression occurs within the cochlea as a consequence of the inherent nonlinear compression of dynamic range, and without it 2TS would vanish ([Ruggero et al., 1992](#)). [Rhode \(2007\)](#) investigated BM 2TS in 16 chinchilla cochleae in the mid-frequency range (5.3 to 9.1 kHz), taking care to ensure the cochleae remained sensitive. [Rhode \(2007\)](#) compared iso-suppression frequency functions (1 dB reduction in probe response) with iso-amplitude functions to the probe alone. The suppression functions were much wider in frequency than the single-tone BM iso-amplitude filter functions. If one were to use these to estimate BM tuning, this would imply that the iso-suppression based estimates give a lower Q_{ERB} than the true iso-amplitude tuning curves. [Rhode \(2007\)](#) reported that suppression extends from the lowest suppression frequency used in his study to nearly 1 octave above CF. Thus, suppression was found to extend throughout and beyond the single-tone tuning region. This is the probable reason for the differences in Q_{ERB} estimates in the present study, which would imply that 2TS yields different tuning properties than single frequency cochlear BM tuning. However, it is not clear to what extent BM vibration iso-suppression tuning curves can be compared with SFOAE iso-suppression tuning curves.

In an attempt to understand the link between suppression and BM vibration patterns, [Lineton and Lutman \(2003\)](#) and [Lineton and Wildgoose \(2009\)](#) modeled 2TS SFOAE tuning curves using the quasilinear long-wave model of [Kanis and de Boer \(1994\)](#). [Lineton and Wildgoose's \(2009\)](#) 2TS tuning curve simulations unfortunately failed to accurately model maximal suppression within the well known empirically determined range $1.1 \leq f_s/f_p \leq 1.2$ ([Brass and Kemp, 1993](#); [Keefe et al., 2008](#); [Lineton and Wildgoose, 2009](#)). Thus key features of 2TS of SFOAEs were not captured in this model. In light of the discrepancies between the SFOAE phase-gradient and 2TS estimates for cochlear tuning in the present study, further modeling work is needed to properly explain 2TS and how it relates to cochlear tuning and BM vibration patterns. The commonly used heuristic argument presented here (see Sec. I B) and in the literature ([Keefe et al., 2008](#); [Lineton and Wildgoose, 2009](#)) is clearly insufficient.

V. SUMMARY AND CONCLUSION

This study compared two SFOAE measures of sharpness-of-tuning Q_{ERB} , based on SFOAE latency and 2TS tuning curves, recorded in the same subjects. The sharpness-of-tuning curves were found to be frequency dependent with sharper tuning from apex to the base of the cochlea. The 2TS Q_{ERB} estimates were found to have a weaker (though statistically significant) frequency dependence than the phase-gradient delay estimates. The key result, however, was that the absolute differences between the two tuning estimates were very large and statistically significant. The Q_{ERB} estimates from both methods were largely in agreement with historical studies.

A major difference in phase-gradient delay estimates existed in the literature between [Shera et al. \(2002\)](#) and [Schairer et al. \(2006\)](#). The present study suggests that these differences were largely due to different post-processing strategies used. Taking a pragmatic view, the results of the present study were obtained using an independent post-processing strategy to that of [Shera et al. \(2002\)](#) and similar results were obtained. Thus, this independent verification would tend to support the results of [Shera et al. \(2002\)](#) over [Schairer et al. \(2006\)](#).

It is obvious and implicit that 2TS estimates of cochlear tuning are strongly affected by suppression in the cochlea. It is argued here that the phase-gradient delay estimates are not similarly affected. Historical studies have shown that SFOAE phase is largely independent of the suppressor frequency used. Therefore, the SFOAE is generated on the BM in the region of tonotopic place, regardless of whether the suppression method is used. An important study by [Rhode \(2007\)](#), in sensitive *in vivo* chinchilla cochleae, pointed to the fact that 2TS tuning curves are much broader in frequency than pure tone BM vibration patterns. This has been used to argue for the differences between the two measures of Q_{ERB} in the present study. Thus, 2TS estimates of cochlear tuning may represent the underlying mechanics of the suppression mechanism, and not necessarily cochlear tuning. However, it is not clear to what extent BM vibration iso-suppression tuning curves can be compared with SFOAE iso-suppression tuning curves, thus highlighting the need for more experimental and modeling verification. The present study concludes that the phase-gradient delay estimate is the most likely one to reflect BM tuning.

ACKNOWLEDGMENTS

The authors wish to thank Christopher Shera and the two anonymous reviewers for their constructive comments.

¹This should not be confused with the 2TS Q_{ERB} measure. The suppression method has been shown by [Kalluri and Shera \(2007a\)](#) to be equivalent to other methods of obtaining SFOAEs (e.g., the compression and spectral smoothing methods) at low excitation levels. The magnitudes and phases of the resulting SFOAEs are nearly equivalent regardless of the method used.

²The range was found by log-transforming preliminary recorded data. With a constant frequency range found on a logarithmic scale across the three bands, it was intended to excite probe and suppressor TW, separated by a constant physical width on the BM across the bands.

³[Lineton and Wildgoose \(2009\)](#) applied a frequency transformation to compensate for the approximate inverse relationship between TW delays and frequency for an approximate scaling-symmetric cochlea ([Shera et al.,](#)

- 2002; Shera and Guinan, 2003). This had the effect of linearizing the SFOAE phase curve in this region. However, Lineton and Wildgoose (2009) found that this transformation made negligible changes to their latency estimates and, thus, this step was neglected here.
- Brass, D., and Kemp, D. (1991). "Time-domain observation of otoacoustic emissions during constant tone stimulation," *J. Acoust. Soc. Am.* **90**, 2415–2427.
- Brass, D., and Kemp, D. (1993). "Suppression of stimulus frequency otoacoustic emissions," *J. Acoust. Soc. Am.* **93**, 920–939.
- Choi, Y.-S., Lee, S.-Y., Parham, K. Neely, S., and Kim, D. (2008). "Stimulus-frequency otoacoustic emission: measurements in humans and simulations with an active cochlear model," *J. Acoust. Soc. Am.* **123**, 2651–2669.
- Glasberg, B., and Moore, B. (1990). "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.* **47**, 103–138.
- Guinan, S. Jr., Backus, B. Lilaonitkul, W., and Aharonson, V. (2003). "Medial olivocochlear efferent reflex in humans: Otoacoustic emission (oae) measurement issues and the advantages of stimulus frequency oaes," *J. Assoc. Res. Otolaryngol.* **4**, 521–540.
- Harte, J., Pigasse, G., and Dau, T. (2009). "Comparison of cochlear delay estimates using otoacoustic emissions and auditory brainstem responses," *J. Acoust. Soc. Am.* **126**, 1291–1301.
- Kalluri, R., and Shera, C. (2001). "Distortion-product source unmixing: A test of the two-mechanism model for DPOAE generation," *J. Acoust. Soc. Am.* **109**, 622–637.
- Kalluri, R., and Shera, C. (2007a). "Comparing stimulus-frequency otoacoustic emissions measured by compression, suppression, and spectral smoothing," *J. Acoust. Soc. Am.* **122**, 3562–3575.
- Kalluri, R., and Shera, C. (2007b). "Near equivalence of human click-evoked and stimulus-frequency otoacoustic emissions," *J. Acoust. Soc. Am.* **121**, 2097–2110.
- Kanis, L., and de Boer, E. (1994). "Two-tone suppression in a locally active nonlinear model of the cochlea," *J. Acoust. Soc. Am.* **96**, 2156–2165.
- Keefe, D., Ellison, J., Fitzpatrick, D., and Gorga, M. (2008). "Two-tone suppression of stimulus frequency otoacoustic emissions," *J. Acoust. Soc. Am.* **123**, 1479–1494.
- Kemp, D. (2008). "Otoacoustic emissions: concepts and origins," in *Active Processes and Otoacoustic Emissions*, Springer Handbook of Auditory Research **30**, 1–38.
- Kemp, D., and Chum, R. (1980). "Properties of the generator of stimulated acoustic emissions," *Hear. Res.* **2**, 213–232.
- Kim, D., Dorn, P., Neely, S., and Gorga, M. (2001). "Adaptation of distortion product otoacoustic emission in humans," *J. Assoc. Res. Otolaryngol.* **2**, 31–40.
- Knight, R., and Kemp, D. (2001). "Wave and place fixed DPOAE maps of the human ear," *J. Acoust. Soc. Am.* **109**, 1513.
- Lineton, B., and Lutman, M. (2003). "Modelling the effect of suppression on the periodicity of stimulus frequency otoacoustic emissions," *J. Acoust. Soc. Am.* **114**, 859–870.
- Lineton, B., and Wildgoose, C. (2009). "Comparing two proposed measures of cochlear mechanical filter bandwidth based on stimulus frequency otoacoustic emissions," *J. Acoust. Soc. Am.* **125**, 1558–1566.
- Moleti, A., and Sisto, R. (2008). "Comparison between otoacoustic and auditory brainstem response latencies supports slow backward propagation of otoacoustic emissions," *J. Acoust. Soc. Am.* **123**, 1495–1503.
- Neely, S., Norton, S., Gorga, M., and Jesteadt, W. (1988). "Latency of auditory brain-stem responses and otoacoustic emissions using tone-burst stimuli," *J. Acoust. Soc. Am.* **83**, 652–656.
- Oxenham, A., and Shera, C. (2003). "Estimates of human cochlear tuning at low levels using forward and simultaneous masking," *J. Assoc. Res. Otolaryngol.* **4**, 541–554.
- Press, W., Teukolsky, S., Vetterling, W., and Flannery, P. (1986). *Numerical Recipes* (Cambridge University Press, New York), pp. 229–232.
- Ren, T. (2004). "Reverse propagation of sound in the gerbil cochlea," *Nat. Neurosci.* **7**, 333–334.
- Rhode, W. (2007). "Mutual suppression in the 6 kHz region of sensitive chinchilla cochleae," *J. Acoust. Soc. Am.* **121**, 2805–2818.
- Rodriguez, J., Neely, S., Patra, H., Kopun, J., Jesteadt, W., Tan, H., and Gorga, M. P. (2010). "The role of suppression in psychophysical tone-on-tone masking," *J. Acoust. Soc. Am.* **127**, 361–369.
- Ruggero, M., Rich, N., Recio, A., Narayan, S., and Robles, L. (1997). "Basilar-membrane responses to tones at the base of the chinchilla cochlea," *J. Acoust. Soc. Am.* **101**, 2151.
- Ruggero, M., Robles, L., and Rich, N. (1992). "Two-tone suppression in the basilar membrane of the cochlea: Mechanical basis of auditory-nerve rate suppression," *J. Neurophysiol.* **68**, 1087–1099.
- Ruggero, M., and Temchin, A. (2007). "Similarity of traveling-wave delays in the hearing organs of humans and other tetrapods," *J. Assoc. Res. Otolaryngol.* **8**, 153–166.
- Schairer, K., Ellison, J., Fitzpatrick, D., and Keefe, D. (2006). "Use of stimulus-frequency otoacoustic emission latency and level to investigate cochlear mechanics in human ears," *J. Acoust. Soc. Am.* **120**, 901–914.
- Schairer, K., Fitzpatrick, D., and Keefe, D. (2003). "Input-output functions for stimulus-frequency otoacoustic emissions in normal-hearing adult ears," *J. Acoust. Soc. Am.* **114**, 944–966.
- Schairer, K., and Keefe, D. (2005). "Simultaneous recording of stimulus-frequency and distortion-product otoacoustic emission input-output functions in human ears," *J. Acoust. Soc. Am.* **117**, 818–832.
- Shera, C., Guinan, J., and Oxenham, A. (2002). "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," *Proc. Natl. Acad. Sci.* **99**, 3318–3323.
- Shera, C., and Guinan, J. Jr., (1999). "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs," *J. Acoust. Soc. Am.* **105**, 782–798.
- Shera, C., and Guinan, J. Jr., (2003). "Stimulus-frequency-emission group delay: A test of coherent reflection filtering and a window on cochlear tuning," *J. Acoust. Soc. Am.* **113**, 2762–2772.
- Shera, C., and Guinan, J. Jr., (2008). "Mechanisms of mammalian otoacoustic emission," in *Active Processes and Otoacoustic Emissions*, Springer Handbook of Auditory Research **30**, 305.
- Shera, C., Guinan, J. Jr., and Oxenham, A. (2010). "Otoacoustic estimation of cochlear tuning: Validation in the chinchilla," *J. Assoc. Res. Otolaryngol.* **11**, 343–365.
- Siegel, J. H., Cerka, A. J., Recio-Spinoso, A., Temchin, A. N., Van Dijk, P., and Ruggero, M. A. (2005). "Delays of stimulus-frequency otoacoustic emissions and cochlear vibrations contradict the theory of coherent reflection filtering," *J. Acoust. Soc. Am.* **118**, 2434–2443.
- Zweig, G. (1976). "Basilar membrane motion," in *Cold Spring Harbor Symposium on Quantitative Biology*, 40, 619 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).